Predictability Assessment and Improving Ensemble Forecasts

Steven L. Mullen, PI
University of Arizona
Department of Atmospheric Sciences
Tucson AZ 85721
TEL (520) 621-6842
FAX (520) 621-6833
E-MAIL mullen@atmo.arizona.edu

Mary M. Poulton, Co-I
University of Arizona
Department of Mining and Geological Engineering
Tucson AZ 85721
TEL (520) 621-8391
FAX (520) 621-8330
E-MAIL mary@mge.arizona.edu

AWARD NUMBER: N00014-99-1-0181

PROJECT GOALS AND OBJECTIVES

The PI and Co-I are examining mesoscale predictability with the goal of improving the utility of ensemble forecasts at ranges of 12 hours to 10 days. Our research addresses the issues of initial condition uncertainty (ICU) for mesoscale analyses, calibration of output from ensemble prediction systems (EPSs) by artificial neural networks, and predictability estimates for precipitation and processes that strongly influence precipitation. The PI also serves as Chief Scientist to Dr. Scott Sandgathe for ONR initiative on Predictability in the Atmosphere and Ocean.

DOCUMENTATION OF ANALYSIS UNCERTAINTY

There is still debate on how to generate initial perturbations for medium-range EPS's with global models. Research on perturbation design for mesoscale limited-area models is, at best, in its infancy. Whatever strategy is employed (dynamic or statistical), it is clear that initial perturbations must be properly constrained by our best estimate of analysis uncertainty.

Several approaches come readily to mind. A promising approach is the integration of the data assimilation system (DAS) and the EPS, the so-called extended Kalman filter (Houtekamer and Mitchell 1998). Another approach is Observing System Simulation Experiments (OSSEs) with a 3DVAR or 4DVAR scheme, which is perhaps the best *proven* way to address the problem. A third method is a thorough documentation of analysis-analysis differences from different analysis-forecast systems. This approach defines a “component" of the analysis uncertainty. Although this methodology is not as theoretically appealing as an integration of the DAS and EPS or as comprehensive as OSSE experiments, it is currently tractable, very economical, and useful guidance can be quickly obtained.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **30 SEP 1999**	2. REPORT TYPE	3. DATES COVERED **00-00-1999 to 00-00-1999**

4. TITLE AND SUBTITLE **Predictability Assessment and Improving Ensemble Forecasts**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Arizona,Department of Atmospheric Sciences,Tucson,AZ,85721**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a REPORT **unclassified**	b ABSTRACT **unclassified**	c THIS PAGE **unclassified**	**Same as Report (SAR)**	**6**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

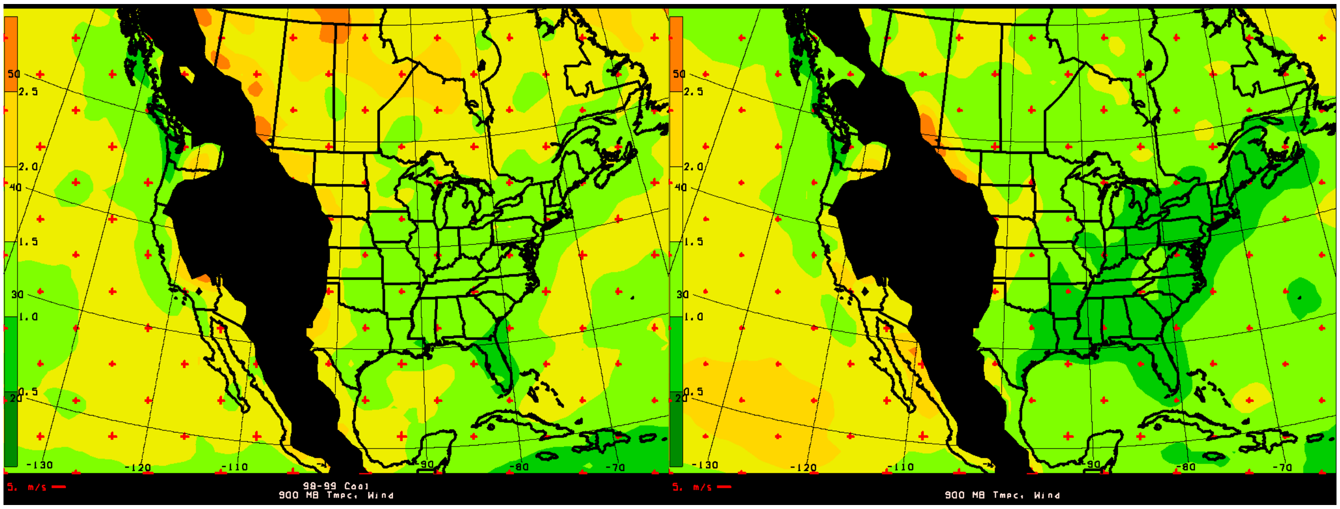

Figure 1: Standard deviation of differences between the ETA and NGM initial analyses for 1998/99 cool season (left) and 1999 warm season (right) for temperature (color fill) and the components of the 900 mb vector wind (line segments). Color fill denotes the ratio of variance of analysis differences to the variance of the ETA analysis: values < 0.5 (green), values between 0.5 and 1.0 (yellow), values > 1.0 (red). Black shading masks areas where model fields are below the ground.

The PI has been comparing differences between NCEP LAM analyses from the NGM (Hoke et al. 1989) and ETA (Rodgers et al. 1995) model. It is important to consider LAM fields since scales not resolved by global analyses are presumably analyzed with greater certainty over the data rich, North American continent. More than one-year of twice-daily analyses has been processed. Figure 1 shows the standard deviation of differences between the ETA and NGM analyses 1998/99 winter and 1999 summer. Plots are shown for the 900 mb temperature and the 900 mb wind components. Note that the oceanic regions and the slopes of mountains are characterized by large absolute uncertainty, consistent with the conventional notion that maritime analyses and the surface layer are more uncertain than the free atmosphere over the continent. There are regions of strong seasonality. Summertime uncertainty is relatively small over the East U.S. but large over West Mexico and the eastern Tropical Pacific. In FY00, the PI will be computing 2D spectra of the difference fields to document the scale dependency of analysis uncertainty.

NEURAL NETWORK POST-PROCESSING OF ENSEMBLE FORECAST PRODUCTS

Because forecast fields produced by any NWP model always contain errors due to model deficiencies (e.g. lack of resolution, inadequate parameterizations, truncation error, etc), raw model output is often statistically post-processed to mitigate their impact. Post-processing also provides a way to relate model output fields to weather elements not explicitly forecast by the NWP model (e.g. visibility, probability of thunder).

There are many viable ways to generate statistical forecasts and calibrate NWP model (e.g. Marzban and Stumpf 1998). The technique currently in use at NCEP (e.g. Carter et al. 1989) is Model Output Statistics (MOS). MOS is based on multiple linear regression (MLR) and typically reduces the error variance by 20%. To yield regression equations, MOS requires a data set of several seasons for training that is averaged over a broad geographic region. The need for such long training data sets is a major shortcoming MLR.

The calibration of EPS output presents even greater challenges than deterministic forecasts because of increased dimensionality of the output. Recent results indicate that strategies besides MLR can be employed for EPS output with equal success (e.g. Hamill and Colucci 1998, Eckel and Walter 1998), methods that correct for biases and other systematic errors (e.g. under dispersion, conditional biases) and require much shorter training periods than MOS.

The mature fields of digit signal processing and artificial intelligence offer other possible paradigms to explore. Consider artificial neural networks (ANNs), computer algorithms designed for nonlinear optimization. Impressive applications of ANNs to remote sensing, atmospheric diagnosis and weather prediction are beginning to appear. A plethora of applications, mostly diagnostic or prognostic based on extrapolation from observations and analyses, recently appeared in the preprint volume of AMS First Conference on Artificial Intelligence (1998); an excellent overview of ANN applications to operational forecasting is given by Christopherson (1998) in the said volume.

The calibration NWP output by ANNs, though limited, are just as impressive (Eckert et al 1996, Hall et al. 1999). The PI's are extending ANN processing to QPF output from the pilot ETA/RSM ensemble data set that was also used by Hamill and Colucci (1998). We find that calibration of 12-24 h rainfall totals at stations with a back propagating ANN, trained with just 12 independent case days, was competitive with MOS. Results obtained under ONR support are shown in Fig. 2. Note that NET markedly improves the RAW ensemble and even shows higher skill than MOS for thresholds up to 1.00". These results will reported in a paper (Mullen et al. 1999, in preparation) that will submitted by the end of the calendar year.

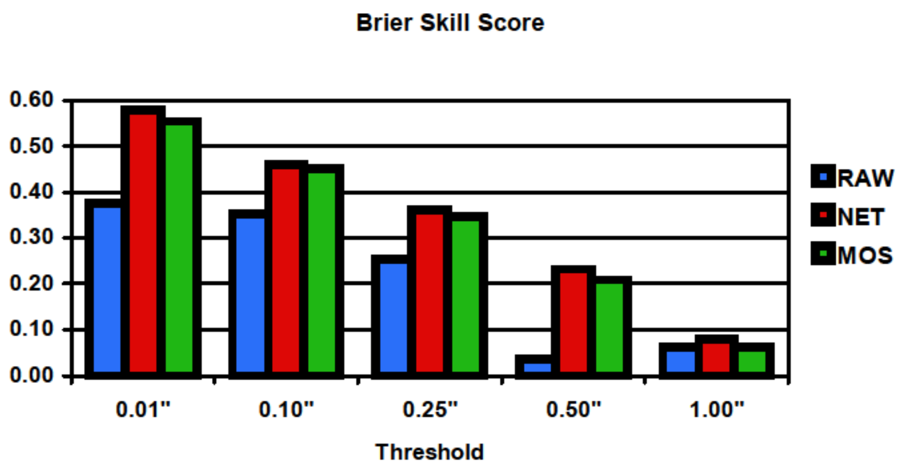

Figure 2. Brier skill scores for various thresholds. Shown are results for the RAW ensemble (blue), NET processing of the ensemble (red) and MOS (green). Skill measured relative to sample climatology. From Mullen et al (1999, in preparation).

Under ONR support in FY00, the PI and Co-I will further examine the utility of ANN processing of EPS output. We will 1) begin processing of precipitation forecasts from the ECMWF EPS and 2) examine training in one geographic region and testing in another with the ECMWF data. We also hope to 1) start screening of predictors and predictands other than precipitation and 2) begin accumulating output from a mesoscale LAM if DURIP funding is awarded. We anticipate that ANN processing of the global and mesoscale LAM EPS output will yield improvements in accuracy comparable to those reported by Hall et al. (1999).

ESTIMATES OF PREDICTABILITY LIMITS FOR PRECIPITATION

The PI and collaborators at NSSL/NOAA and ECMWF are estimating predictability limits for QPF and related fields. The curve for one-hour accumulations of MM5 precipitation on a 32-km grid becomes flat by about 6 h (Fig. 3), consistent with nonlinear saturation and loss of probabilistic skill. The 6-h totals saturate by 12 h and 12 h totals by 24 h. Convective available potential energy (CAPE) appears to saturate before 6 h (not shown). The 80-km NCEP ensemble does produces skillful forecasts of 24-h accumulations out to 2 days (not shown). Cost-loss analysis indicates that the NCEP ensemble has higher value than the deterministic meso-ETA for thresholds up to 1.00" per day (not shown), the highest threshold examined. These results are described in a paper (Wandishin et al 1999) under review.

Results for the ECMWF EPS are shown in Fig. 4. The EPS forecasts are more skillful during the winter than the summer. The EPS produces skillful forecasts to past 10 days for a threshold of 1 mm in both seasons. Accuracy decreases as the threshold increases, until forecasts of 50 mm are not skillful at 1 day during summer. The longer range of skill for the ECMWF forecasts relative to the LAM foreacasts is due to the longer accumulation period and coarser resolution of the model. These results are described in a paper (Mullen and Buizza) under review.

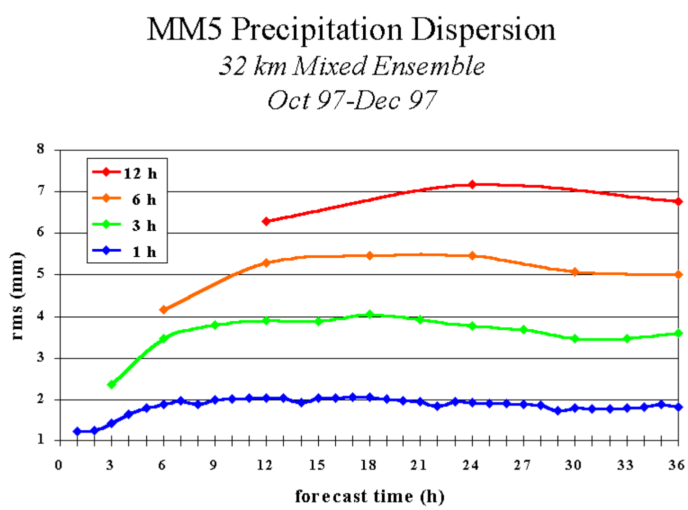

Figure 3: Dispersion of precipitation for various accumulation periods for a mixed MM5 ensemble. From Wandishin et al (1999 submitted).

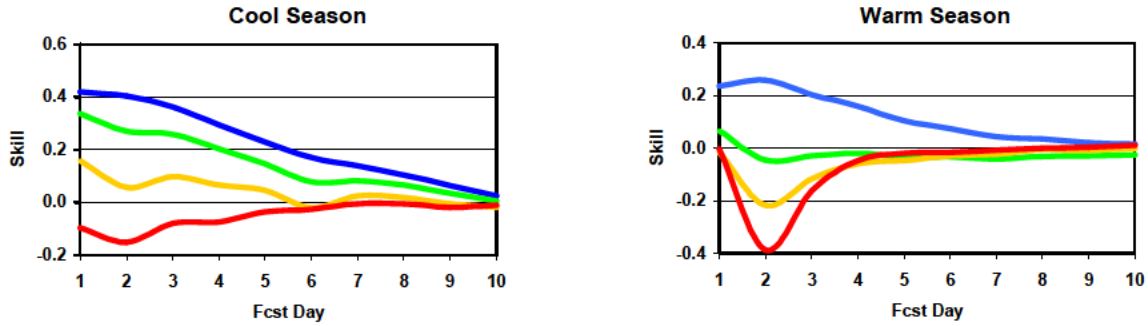

Figure 4: Brier skill scores, relative to long-term climatology, for the cool and warm seasons. Thresholds are 1 mm (blue), 10 mm (green), 20 mm (gold), and 50 mm (red). Note different scales for the ordinates. From Mullen and Buizza (1999, submitted).

Under ONR support in FY00, the PI will be examining optimal configurations for ensemble forecast systems. We also hope to explore predictability limits and high membership ensembles *O*(100) from daily forecasts with a mesoscale LAM (~10 km grid) if pending DURIP funding is awarded.

REFERENCES

Carter, G.M., J.P. Dallavalle, and H.R. Glahn, 1989: Statistical forecasts based on the National Meteorological Center's numerical weather prediction system. Wea. Forecasting, 4, 401-412.

Christopherson, D., 1998: Artificial intelligence in the weather forecast office: one forecaster's view. Preprints, 1st Conference on Artificial Intelligence. 11-16 January 1998. Phoenix, Arizona. American Meteorology Society. 136-143.

Hamill, T.M. and S.J. Colucci, 1998: Verification of Eta/RSM ensemble probabilistic precipitation forecasts. Mon. Wea. Rev., 126, 711-724.

Hoke, J. E., N. A. Phillips, G. J. DiMego, J. J. Tuccillo, and J. G. Sela, 1989: The Regional Analysis and Forecast system of the National Meteorological Center. Wea. Forecasting, 4, 323-334.

Houtekamer, P.L., and H.L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. Mon. Wea. Rev., 126, 796-811 .

Marzban, C., and G. Stumpf, 1998: A neural network for damaging wind prediction. Wea. Forecasting, 13, 151-163.

Rogers, E., T. Black, D. Deaven, G. Dimego, Q. Zhao, Y. Lin, N. Junker, M. Baldwin, 1995: Changes to the NMC Operational ETA Model Analysis/Forecast System. Technical Procedures Bulletin 423, NWS Office of Meteorology, Silver Spring, MD. 51 pp.

PUBLICATIONS (SUBMITTED AND UNDER REVIEW)

Mullen, S.L. and R. Buizza, 1999: Quantitative precipitation forecasts over the United States by the ECMWF ensemble prediction system. Mon. Wea. Rev., 1999, submitted.

Mullen, S.L., M. M. Poulton, H. E. Brooks, T. M. Hamill, 1999: Calibration of ensemble precipitation forecasts by an artificial neural network. Wea. Forecasting, to be submitted by late 1999.

Wandishin, M.S., S.L. Mullen, D.J. Stensrud, and H.E. Brooks, 1999: Evaluation of a short-range multi-model ensemble. Mon. Wea. Rev., submitted.

IN-HOUSE/OUT-OF-HOUSE RATIOS

All research is 100% out-of-house.